

0590  
0702

#5 OIPE

## RAW SEQUENCE LISTING

DATE: 07/09/2002

PATENT APPLICATION: US/10/074,328

TIME: 15:03:42

Input Set : N:\CrF3\RULE60\10074328.raw

Output Set: N:\CRF3\07092002\J074328.raw

## SEQUENCE LISTING

ENTERED

3 (1) GENERAL INFORMATION:

5 (i) APPLICANT: BROW, MARY ANN D.

6 GROTELUESCHEN HALL, JEFF S.

7 LYAMICHEV, VICTOR

8 OLIVE, DAVID M.

9 PRUDENT, JAMES R.

11 (ii) TITLE OF INVENTION: DETECTION OF NUCLEIC ACID SEQUENCES BY

12 INVADER-DIRECTED CLEAVAGE

14 (iii) NUMBER OF SEQUENCES: 48

16 (iv) CORRESPONDENCE ADDRESS:

17 (A) ADDRESSEE: MEDLEN & CARROLL

18 (B) STREET: 220 MONTGOMERY STREET, SUITE 2200

19 (C) CITY: SAN FRANCISCO

20 (D) STATE: CALIFORNIA

21 (E) COUNTRY: UNITED STATES OF AMERICA

22 (F) ZIP: 94104

24 (v) COMPUTER READABLE FORM:

25 (A) MEDIUM TYPE: Floppy disk

26 (B) COMPUTER: IBM PC compatible

27 (C) OPERATING SYSTEM: PC-DOS/MS-DOS

28 (D) SOFTWARE: PatentIn Release #1.0, Version #1.25

30 (vi) CURRENT APPLICATION DATA:

C--> 31 (A) APPLICATION NUMBER: US/10/074,328

C--> 32 (B) FILING DATE: 12-Feb-2002

33 (C) CLASSIFICATION:

36 (vii) PRIOR APPLICATION DATA:

37 (A) APPLICATION NUMBER: US/08/599,491

38 (B) FILING DATE: 23-JAN-1996

40 (viii) ATTORNEY/AGENT INFORMATION:

41 (A) NAME: INGOLIA, DIANE E.

42 (B) REGISTRATION NUMBER: P-40,027

43 (C) REFERENCE/DOCKET NUMBER: FORS-01802

45 (ix) TELECOMMUNICATION INFORMATION:

46 (A) TELEPHONE: (415) 705-8410

47 (B) TELEFAX: (415) 397-8338

50 (2) INFORMATION FOR SEQ ID NO: 1:

52 (i) SEQUENCE CHARACTERISTICS:

53 (A) LENGTH: 2506 base pairs

54 (B) TYPE: nucleic acid

55 (C) STRANDEDNESS: double

56 (D) TOPOLOGY: linear

58 (ii) MOLECULE TYPE: DNA (genomic)

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/074,328

DATE: 07/09/2002

TIME: 15:03:42

Input Set : N:\Crif3\RULE60\10074328.raw

Output Set: N:\CRF3\07092002\J074328.raw

```

62      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:
64      ATGAGGGGGA TGCTGCCCCCT CTTTGAGCCC AAGGGCCGGG TCCTCCTGGT GGACGGCCAC      60
66      CACCTGGCCT ACCGCACCTT CCACGCCCTG AAGGGCCTCA CCACCAGCCG GGGGGAGCCG      120
68      GTGCAGGCGG TCTACGGCTT CGCCAAGAGC CTCCTCAAGG CCCTCAAGGA GGACGGGGAC      180
70      GCGGTGATCG TGGTCTTTGA CGCCAAGGCC CCCTCCTTCC GCCACGAGGC CTACGGGGGG      240
72      TACAAGGCGG GCGGGGCCCC CACGCCGGAG GACTTTCCCC GGCAACTCGC CCTCATCAAG      300
74      GAGCTGGTGG ACCTCCTGGG GCTGGCGCGC CTCGAGGTCC CGGGCTACGA GGCGGACGAC      360
76      GTCCTGGCCA GCCTGGCCAA GAAGGCGGAA AAGGAGGGCT ACGAGGTCCG CATCCTCACC      420
78      GCCGACAAAG ACCTTTACCA GCTCCTTTCC GACCGCATCC ACGTCCTCCA CCCCAGAGGG      480
80      TACCTCATCA CCCC GGCCCTG GCTTTGGGAA AAGTACGGCC TGAGGCCCGA CCAGTGGGCC      540
82      GACTACCGGG CCCTGACCGG GGACGAGTCC GACAACCTTC CCGGGGTCAA GGGCATCGGG      600
84      GAGAAGACGG CGAGGAAGCT TCTGGAGGAG TGGGGGAGCC TGGAAGCCCT CCTCAAGAAC      660
86      CTGGACCGGC TGAAGCCCGC CATCCGGGAG AAGATCCTGG CCCACATGGA CGATCTGAAG      720
88      CTCTCCTGGG ACCTGGCCAA GGTGCGCACC GACCTGCCCC TGGAGGTGGA CTTGCCCAAA      780
90      AGGCGGGAGC CCGACCGGGA GAGGCTTAGG GCCTTTCTGG AGAGGCTTGA GTTTGGCAGC      840
92      CTCCTCCACG AGTTCGGCCT TCTGGAAGC CCAAGGCC TGGAGGAGGC CCCCTGGCCC      900
94      CCGCCGGAAG GGGCCTTCGT GGGCTTTGTG CTTTCCCGCA AGGAGCCCAT GTGGGCCGAT      960
96      CTTCTGGCCC TGGCCGCCGC CAGGGGGGGC CGGGTCCACC GGGCCCCCGA GCCTTATAAA      1020
98      GCCCTCAGG ACCTGAAGGA GGCGCGGGG CTTCTCGCCA AAGACCTGAG CGTTCTGGCC      1080
100     CTGAGGGAAG GCCTTGGCCT CCCGCCGGG GACGACCCCA TGCTCCTCGC CTACCTCCTG      1140
102     GACCCTTCCA ACACACCCC CGAGGGCGTG GCCCGCGCT ACGGCGGGGA GTGGACGGAG      1200
104     GAGGCGGGG AGCGGGCCGC CCTTTCCGAG AGGCTCTTCG CCAACCTGTG GGGGAGGCTT      1260
106     GAGGGGGAGG AGAGGCTCCT TTGGCTTTAC CGGGAGGTGG AGAGGCCCTT TTCCGCTGTC      1320
108     CTGGCCACA TGGAGGCAC GGGGGTGCGC CTGGACGTGG CCTATCTCAG GGCCTGTGCC      1380
110     CTGGAGGTGG CCGAGGAGAT CGCCCGCCTC GAGGCGGAGG TCTTCCGCCT GGCCGGCCAC      1440
112     CCCTTCAACC TCAACTCCCG GGACCAGCTG GAAAGGTGCC TCTTTGACGA GCTAGGGCTT      1500
114     CCCGCCATCG GCAAGACGGA GAAGACCGGC AAGCGCTCCA CCAGCGCCGC CGTCCTGGAG      1560
116     GCCCTCCGCG AGGCCACCC CATCGTGGAG AAGATCCTGC AGTACCGGGA GCTACCAAG      1620
118     CTGAAGAGCA CCTACATTGA CCCCTTGCCG GACCTCATCC ACCCAGGAC GGGCCGCCTC      1680
120     CACACCCGCT TCAACCAGAC GGCCACGGCC ACGGCGAGG TAAGTAGCTC CGATCCCAAC      1740
122     CTCCAGAACA TCCCCGTCCG CACCCCGCTT GGGCAGAGGA TCCGCCGGGC CTTTCATCGC      1800
124     GAGGAGGGGT GGCTATTGGT GGCCCTGGAC TATAGCCAGA TAGAGCTCAG GGTGCTGGCC      1860
126     CACCTCTCCG GCGACGAGAA CCTGATCCGG GTCTTCCAGG AGGGGCGGGA CATCCACACG      1920
128     GAGACGCCCA GCTGGATGTT CGGCGTCCCC CGGGAGGCCG TGGACCCCTT GATGCGCCGG      1980
130     GCGGCCAAGA CCATCAACTT CGGGGTCTCT TACGGCATGT CGGCCACCG CCTCTCCAG      2040
132     GAGTAGCCA TCCCTTACGA GGAGGCCAG GCCTTCATTG AGCGCTACTT TCAGAGCTTC      2100
134     CCAAGGTGC GGGCTGGAT TGAGAAGACC CTGGAGGAG GCAGGAGGCG GGGGTACGTG      2160
136     GAGACCCTCT TCGGCCGCC CCGTACGTG CCAGACCTAG AGGCCCGGT GAAGAGCGTG      2220
138     CGGGAGGCGG CCGAGCGCAT GGCCTTCAAC ATGCCCGTCC AGGGCACCGC CGCCGACCTC      2280
140     ATGAAGCTGG CTATGGTGAA GCTCTTCCCC AGGCTGGAGG AAATGGGGGC CAGGATGCTC      2340
142     CTTCAGGTCC ACGACGAGCT GGTCTCTGAG GCCCCAAAAG AGAGGGCGGA GGCCGTGGCC      2400
144     CGGCTGGCCA AGGAGGTCAT GGAGGGGGTG TATCCCCTGG CCGTGCCCTT GGAGGTGGAG      2460
146     GTGGGGATAG GGGAGGACTG GCTCTCCGCC AAGGAGTGAT ACCACC      2506

```

148 (2) INFORMATION FOR SEQ ID NO: 2:

150 (i) SEQUENCE CHARACTERISTICS:

151 (A) LENGTH: 2496 base pairs

152 (B) TYPE: nucleic acid

153 (C) STRANDEDNESS: double

154 (D) TOPOLOGY: linear

## RAW SEQUENCE LISTING

DATE: 07/09/2002

PATENT APPLICATION: US/10/074,328

TIME: 15:03:42

Input Set : N:\Crf3\RULE60\10074328.raw

Output Set: N:\CRF3\07092002\J074328.raw

```

156      (ii) MOLECULE TYPE: DNA (genomic)
160      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:
162 ATGGCGATGC TTCCCCTCTT TGAGCCCAAA GGCCGCGTGC TCCTGGTGA CGGCCACCAC      60
164 CTGGCCTACC GCACCTTCTT TGCCCTCAAG GGCCCTACCA CCAGCCGCGG CGAACCCGTT      120
166 CAGGCGGTCT ACGGCTTTCG CAAAAGCCTC CTCAAGGCCC TGAAGGAGGA CGGGGACGTG      180
168 GTGGTGGTGG TCTTTGACGC CAAGGCCCCC TCCTTCCGCC ACGAGGCCTA CGAGGCCTAC      240
170 AAGGCGGGCC GGGCCCCCAC CCCGGAGGAC TTTCCCCGGC AGCTGGCCCT CATCAAGGAG      300
172 TTGGTGGACC TCCTAGGCCT TGTGCGGCTG GAGGTTCCCG GCTTTGAGGC GGACGACGTG      360
174 CTGGCCACCC TGGCCAAGCG GCGGAAAAG GAGGGGTACG AGGTGCGCAT CCTCACTGCC      420
176 GACCGCGACC TCTACCAGCT CCTTTCGGAG CGCATCGCCA TCCTCCACCC TGAGGGGTAC      480
178 CTGATCACCC CGGCGTGGCT TTACGAGAAG TACGGCCTGC GCCCGGAGCA GTGGGTGGAC      540
180 TACCGGGCCC TGGCGGGGA CCCCTCGGAT AACATCCCCG GGGTGAAGGG CATCGGGGAG      600
182 AAGACCGCCC AGAGGCTCAT CCGCGAGTGG GGGAGCCTGG AAAACCTCTT CCAGCACCTG      660
184 GACCAGGTGA AGCCCTCCTT GCGGGAGAAG CTCCAGGCGG GCATGGAGGC CCTGGCCCTT      720
186 TCCCGGAAGC TTTCCAGGT GCACACTGAC CTGCCCCTGG AGGTGGACTT CGGGAGGCGC      780
188 CGCACACCCA ACCTGGAGGG TCTGCGGGCT TTTTGGAGC GGTGGAGTT TGAAGCCCTC      840
190 CTCCACGAGT TCGGCCTCCT GGAGGGGCCG AAGCGGCAG AGGAGGCCCC CTGGCCCCCT      900
192 CCGGAAGGGG CTTTTTGGG CTTTTCTTTT TCCCGTCCC AGCCCATGTG GGCCGAGCTT      960
194 CTGGCCCTGG CTGGGGCGTG GGAGGGGCGC CTCCATCGGG CACAAGACCC CCTTAGGGGC      1020
196 CTGAGGGACC TTAAGGGGT GCGGGGAATC CTGGCCAAGG ACCTGGCGGT TTTGGCCCTG      1080
198 CGGGAGGGCC TGGACCTCTT CCCAGAGGAC GACCCCATGC TCCTGGCCTA CCTTCTGGAC      1140
200 CCCTCCAACA CCACCCCTGA GGGGGTGGCC CGGCGTTACG GGGGGGAGTG GACGGAGGAT      1200
202 GCGGGGGAGA GGGCCCTCCT GGCCGAGCGC CTCTTCCAGA CCCTAAAGGA GCGCCTTAAG      1260
204 GGAGAAGAAC GCCTGCTTTG GCTTTACGAG GAGGTGGAGA AGCCGCTTTC CCGGGTGTG      1320
206 GCCCGGATGG AGGCCACGGG GGTCCGGCTG GACGTGGCCT ACCTCCAGGC CCTCTCCCTG      1380
208 GAGGTGGAGG CGGAGGTGCG CCAGCTGGAG GAGGAGTCT TCCGCCTGGC CGGCCACCCC      1440
210 TTCAACCTCA ACTCCGCGA CCAGCTGGAG CGGGTGCTCT TTGACGAGCT GGGCCTGCCT      1500
212 GCCATCGGCA AGACGGAGAA GACGGGAAA CGCTCCACCA GCGCTGCCGT GCTGGAGGCC      1560
214 CTGCGAGAGG CCCACCCAT CGTGGACCGC ATCCTGCAGT ACCGGGAGCT CACCAAGCTC      1620
216 AAGAACACCT ACATAGACCC CCTGCCC GCC CTGGTCCACC CCAAGACCG CCGGCTCCAC      1680
218 ACCCGCTTCA ACCAGACGGC CACCGCCACG GGCAGGCTTT CCAGCTCCGA CCCCACCTG      1740
220 CAGAACATCC CCGTGCGCAC CCTCTGGGC CAGCGCATCC GCCGAGCCTT CGTGCCGAG      1800
222 GAGGGCTGGG TGCTGGTGGT CTTGGACTAC AGCCAGATTG AGCTTCGGGT CCTGGCCAC      1860
224 CTCTCCGGG ACGAGAACCT GATCCGGGT TTTCAGGAG GGAGGGACAT CCACACCCAG      1920
226 ACCGCCAGCT GGATGTTCCG CGTTTCCCC GAAGGGGTAG ACCCTCTGAT GCGCCGGCG      1980
228 GCCAAGACCA TCAACTTCGG GGTGCTCTAC GGCATGTCCG CCCACCGCCT CTCCGGGGAG      2040
230 CTTTCCATCC CCTACGAGGA GGCGGTGGCC TTATTGAGC GCTACTTCCA GAGCTACCCC      2100
232 AAGGTGCGGG CCTGGATTGA GGGGACCCTC GAGGAGGGCC GCCGGCGGGG GTATGTGGAG      2160
234 ACCCTCTTCG GCCGCCGGCG CTATGTGCC GACCTCAACG CCCGGGTGAA GAGCGTGCGC      2220
236 GAGGCGGCG AGCGCATGGC CTTCAACATG CCGGTCCAGG GCACCGCCGC CGACCTCATG      2280
238 AAGCTGGCCA TGGTGC GGCT TTTCCCCCG CTTCAGGAAC TGGGGGCGAG GATGCTTTTG      2340
240 CAGGTGCACG ACGAGCTGGT CCTCGAGGCC CCAAGGACC GGGCGGAGAG GGTAGCCGCT      2400
242 TTGGCCAAGG AGGTCATGGA GGGGGTCTGG CCCCTGCAGG TGCCCTGGA GGTGGAGGTG      2460
244 GGCCTGGGGG AGGACTGGCT CTCCGCAAG GAGTAG      2496
246 (2) INFORMATION FOR SEQ ID NO: 3:
248      (i) SEQUENCE CHARACTERISTICS:
249          (A) LENGTH: 2504 base pairs
250          (B) TYPE: nucleic acid
251          (C) STRANDEDNESS: double

```

## RAW SEQUENCE LISTING

DATE: 07/09/2002

PATENT APPLICATION: US/10/074,328

TIME: 15:03:42

Input Set : N:\Crif3\RULE60\10074328.raw

Output Set: N:\CRF3\07092002\J074328.raw

```

252         (D) TOPOLOGY: linear
254         (ii) MOLECULE TYPE: DNA (genomic)
258         (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:
260 ATGGAGGCGA TGCTTCCGCT CTTTGAACCC AAAGGCCGGG TCCTCCTGGT GGACGGCCAC      60
262 CACCTGGCCT ACCGCACCTT CTTTCGCCCTG AAGGGCCTCA CCACGAGCCG GGGCGAACC      120
264 GTGCAGGCGG TCTACGGCTT CGCCAAGAGC CTCCTCAAGG CCCTGAAGGA GGACGGGTAC      180
266 AAGGCCGTCT TCGTGGTCTT TGACGCCAAG GCCCCCTCCT TCCGCCACGA GGCCTACGAG      240
268 GCCTACAAGG CGGGGAGGGC CCCGACCCCC GAGGACTTCC CCCGGCAGCT CGCCCTCATC      300
270 AAGGAGCTGG TGGACCTCCT GGGGTTTACC CGCCTCGAGG TCCCCGGCTA CGAGGCGGAC      360
272 GACGTTCTCG CCACCCTGGC CAAGAAGGCG GAAAAGGAGG GGTACGAGGT GCGCATCCTC      420
274 ACCGCCGACC GCGACCTCTA CCAACTCGTC TCCGACCGCG TCGCCGTCCT CCACCCCGAG      480
276 GGCCACCTCA TCACCCCGGA GTGGCTTTGG GAGAAGTACG GCCTCAGGCC GGAGCAGTGG      540
278 GTGGACTTCC GCGCCCTCGT GGGGGACCCC TCCGACAACC TCCCCGGGGT CAAGGGCATC      600
280 GGGGAGAAGA CCGCCCTCAA GCTCCTCAAG GAGTGGGGAA GCCTGGA AAA CCTCCTCAAG      660
282 AACCTGGACC GGGTAAAGCC AGAAAACGTC CGGGAGAAGA TCAAGGCCCA CCTGGAAGAC      720
284 CTCAGGCTCT CTTGGAGCT CTCCCGGGTG CGCACCGACC TCCCCCTGGA GGTGGACCTC      780
286 GCCCAGGGGC GGGAGCCCGA CCGGGAGGGG CTTAGGGCCT TCCTGGAGAG GCTGGAGTTC      840
288 GGCAGCCTCC TCCACGAGTT CGGCCTCCTG GAGGCCCCCG CCCCCCTGGA GGAGGCCCCC      900
290 TGGCCCCCGC CGGAAGGGGC CTTCTGTGGC TTCGTCTCTT CCCGCCCGA GCCCATGTGG      960
292 GCGGAGCTTA AAGCCCTGGC CGCCTGCAGG GACGGCCGGG TGCACCGGGC AGCAGACCCC     1020
294 TTGGCGGGGC TAAAGGACCT CAAGGAGGTC CGGGGCCTCC TCGCCAAGGA CCTCGCCGTC     1080
296 TTGGCCTCGA GGGAGGGGCT AGACCTCGTG CCCGGGGACG ACCCCATGCT CCTCGCCTAC     1140
298 CTCCTGGACC CCTCCAACAC CACCCCGAG GGGGTGGCGC GCGCTACGG GGGGGAGTGG     1200
300 ACGGAGGACG CCGCCACCG GGCCCTCCTC TCGGAGAGGC TCCATCGGAA CCTCCTTAAG     1260
302 CGCCTCGAGG GGGAGGAGAA GCTCCTTTGG CTCTACCACG AGGTGGA AAA GCCCCTCTCC     1320
304 CGGGTCCTGG CCCACATGGA GGCCACCGGG GTACGGCTGG ACGTGGCCTA CCTTCAGGCC     1380
306 CTTTCCCTGG AGCTTGCGGA GGAGATCCGC CGCCTCGAGG AGGAGGTCTT CCGCTTGGCG     1440
308 GGCCACCCCT TCAACCTCAA CTCCCGGGAC CAGCTGGAAA GGGTGCTCTT TGACGAGCTT     1500
310 AGGCTTCCCG CTTGGGGGAA GACGCAAAAAG ACAGGCAAGC GCTCCACCAG CGCCGCGGTG     1560
312 CTGGAGGCCC TACGGGAGGC CCACCCCATC GTGGAGAAGA TCCTCCAGCA CCGGGAGCTC     1620
314 ACCAAGCTCA AGAACACCTA CGTGGACCCC CTCCAAGCC TCGTCCACCC GAGGACGGGC     1680
316 CGCCTCCACA CCCGCTTCAA CCAGACGGCC ACGGCCACGG GGAGGCTTAG TAGCTCCGAC     1740
318 CCCAACCTGC AGAACATCCC CGTCCGCACC CCCTTGGGCC AGAGGATCCG CCGGGCCTTC     1800
320 GTGGCCGAGG CGGGTTGGGC GTTGGTGGCC CTGGACTATA GCCAGATAGA GCTCCGCGTC     1860
322 CTCGCCACC TCTCCGGGGA CGAAAACCTG ATCAGGTCTT TCCAGGAGGG GAAGGACATC     1920
324 CACACCAGA CCGCAAGCTG GATGTTCCGC GTCCCCCGG AGGCCGTGGA CCCCCTGATG     1980
326 CGCCGGGCGG CCAAGACGGT GAACTTCGGC GTCCTCTACG GCATGTCCGC CCATAGGCTC     2040
328 TCCCAGGAGC TTGCCATCCC CTACGAGGAG GCGGTGGCCT TTATAGAGGC TACTTCCAAA     2100
330 GCTTCCCCAA GGTGCGGGCC TGGATAGAAA AGACCCTGGA GGAGGGGAGG AAGCGGGGCT     2160
332 ACGTGGAAAC CCTCTTCGGA AGAAGGCGCT ACGTGCCCGA CCTCAACGCC CGGGTGAAGA     2220
334 GCGTCAGGGA GGCCGCGGAG CGCATGGCCT TCAACATGCC CGTCCAGGGC ACCGCCGCCG     2280
336 ACCTCATGAA GCTCGCCATG GTGAAGCTCT TCCCCGCTT CCGGGAGATG GGGGCCCGCA     2340
338 TGCTCCTCCA GGTCCACGAC GAGCTCCTCC TGGAGGCCCC CCAAGCGCGG GCCGAGGAGG     2400
340 TGGCGGCTTT GGCCAAGGAG GCCATGGAGA AGGCCTATCC CCTCGCCGTG CCCCTGGAGG     2460
342 TGGAGGTGGG GATGGGGGAG GACTGGCTTT CCGCAAGGG TTAG                               2504
344 (2) INFORMATION FOR SEQ ID NO: 4:
346         (i) SEQUENCE CHARACTERISTICS:
347             (A) LENGTH: 832 amino acids
348             (B) TYPE: amino acid

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/074,328

DATE: 07/09/2002

TIME: 15:03:42

Input Set : N:\Crf3\RULE60\10074328.raw

Output Set: N:\CRF3\07092002\J074328.raw

```

349         (C) STRANDEDNESS: single
350         (D) TOPOLOGY: linear
352 (ii) MOLECULE TYPE: protein
356 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:
358 Met Arg Gly Met Leu Pro Leu Phe Glu Pro Lys Gly Arg Val Leu Leu
359 1          5          10          15
361 Val Asp Gly His His Leu Ala Tyr Arg Thr Phe His Ala Leu Lys Gly
362          20          25          30
364 Leu Thr Thr Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly Phe Ala
365          35          40          45
367 Lys Ser Leu Leu Lys Ala Leu Lys Glu Asp Gly Asp Ala Val Ile Val
368          50          55          60
370 Val Phe Asp Ala Lys Ala Pro Ser Phe Arg His Glu Ala Tyr Gly Gly
371          65          70          75          80
373 Tyr Lys Ala Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro Arg Gln Leu
374          85          90          95
376 Ala Leu Ile Lys Glu Leu Val Asp Leu Leu Gly Leu Ala Arg Leu Glu
377          100         105         110
379 Val Pro Gly Tyr Glu Ala Asp Asp Val Leu Ala Ser Leu Ala Lys Lys
380          115         120         125
382 Ala Glu Lys Glu Gly Tyr Glu Val Arg Ile Leu Thr Ala Asp Lys Asp
383          130         135         140
385 Leu Tyr Gln Leu Leu Ser Asp Arg Ile His Val Leu His Pro Glu Gly
386          145         150         155         160
388 Tyr Leu Ile Thr Pro Ala Trp Leu Trp Glu Lys Tyr Gly Leu Arg Pro
389          165         170         175
391 Asp Gln Trp Ala Asp Tyr Arg Ala Leu Thr Gly Asp Glu Ser Asp Asn
392          180         185         190
394 Leu Pro Gly Val Lys Gly Ile Gly Glu Lys Thr Ala Arg Lys Leu Leu
395          195         200         205
397 Glu Glu Trp Gly Ser Leu Glu Ala Leu Leu Lys Asn Leu Asp Arg Leu
398          210         215         220
400 Lys Pro Ala Ile Arg Glu Lys Ile Leu Ala His Met Asp Asp Leu Lys
401          225         230         235         240
403 Leu Ser Trp Asp Leu Ala Lys Val Arg Thr Asp Leu Pro Leu Glu Val
404          245         250         255
406 Asp Phe Ala Lys Arg Arg Glu Pro Asp Arg Glu Arg Leu Arg Ala Phe
407          260         265         270
409 Leu Glu Arg Leu Glu Phe Gly Ser Leu Leu His Glu Phe Gly Leu Leu
410          275         280         285
412 Glu Ser Pro Lys Ala Leu Glu Glu Ala Pro Trp Pro Pro Pro Glu Gly
413          290         295         300
415 Ala Phe Val Gly Phe Val Leu Ser Arg Lys Glu Pro Met Trp Ala Asp
416          305         310         315         320
418 Leu Leu Ala Leu Ala Ala Ala Arg Gly Gly Arg Val His Arg Ala Pro
419          325         330         335
421 Glu Pro Tyr Lys Ala Leu Arg Asp Leu Lys Glu Ala Arg Gly Leu Leu
422          340         345         350
424 Ala Lys Asp Leu Ser Val Leu Ala Leu Arg Glu Gly Leu Gly Leu Pro

```

RAW SEQUENCE LISTING ERROR SUMMARY  
PATENT APPLICATION: US/10/074,328

DATE: 07/09/2002  
TIME: 15:03:43

Input Set : N:\Crf3\RULE60\10074328.raw  
Output Set: N:\CRF3\07092002\J074328.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:7; N Pos. 4,5,181,182,190,366,617,628,685,714,722,738,784,1022,1029

Seq#:7; N Pos. 1038,1053,1098,1105,1206,1227,1244,1251,1252,1253,1350,1380

Seq#:7; N Pos. 1497,1530,1569,1572,1641,1653,1655,1770,1812,2319,2346,2396

Seq#:8; Xaa Pos.2,63,109,186,205,209,227,228,233,240,243,244,247,260,290

Seq#:8; Xaa Pos.329,336,340,368,414,417,418,431,551,605,773,794,798,823,833

## VERIFICATION SUMMARY

PATENT APPLICATION: US/10/074,328

DATE: 07/09/2002

TIME: 15:03:43

Input Set : N:\Crf3\RULE60\10074328.raw

Output Set: N:\CRF3\07092002\J074328.raw

L:31 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]  
L:32 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]  
L:970 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8 after pos.:0  
L:979 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8 after pos.:48  
L:988 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8 after pos.:96  
L:1003 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8 after pos.:176  
L:1006 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8 after pos.:192  
L:1009 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8 after pos.:208  
L:1012 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8 after pos.:224  
L:1015 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8 after pos.:240  
L:1018 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8 after pos.:256  
L:1024 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8 after pos.:288  
L:1030 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8 after pos.:320  
L:1033 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8 after pos.:336  
L:1036 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8 after pos.:352  
L:1045 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8 after pos.:400  
L:1048 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8 after pos.:416  
L:1072 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8 after pos.:544  
L:1081 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8 after pos.:592  
L:1114 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8 after pos.:768  
L:1117 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8 after pos.:784  
L:1123 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8 after pos.:816  
L:1126 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8 after pos.:832